A

- DNA in body – pluto and back – 10 B miles
- One cell – 3 metres
- Chromosomes are meaningless. Location means nothing.

Parasitic 53%

Central dogma – FIGURES 2.2, 2.3

Virus, retrovirus, endogenous retrovirus, transposon. -- FIGURE 4.1 & ZIMMER

Battlefield of corpses from viruses wars, explosions/radiations and mass extinctions, parallel biosphere/cell history.

Can tell from their code that these are inactivated virus relics and remnants. Cannot be functional in themselves. So why do we have them?

CONSERVED – Many lineages are over 100 My old, stopped jumping long ago. Mammals keep them around much longer. In flies, half life 12 My; mammals 800 My.

BENEFIT recombination by random inserts, Plague culling. Pool of preadapted mixing modules (transpositions and lateral transfer) to increase evolution, increase adaptability, decrease extinctions, natural genetic engineering maximizes use of it. It's not the junk per se that's beneficial, it's the process that created it and the side effects of it that are beneficial.

EXAMPLE: 100 miRNAs arose from transposons and retrotransposons. -Shapiro

Proportions: some of the coding genes lie within the deactivated virus sequences. Vehicle junk yard.

ENCODE (Encyclopedia of DNA Elements)- 80% functional

(at some point in some cell types, mostly in development)

RNA (highly sensitive to small amounts): Artifact (microarrays versus sequencing); only tested 1%; not conserved; byproduct (Jackson Pollack); noisy (leaky, accidental).
Other assumptions: functional epigenetic modifications, looping, accessible to enzymes,

→if 80% transcribed but 60% is fossils, therefore majority of it must be random byproduct that is inactivated or untranslated! **READ PAGE - STOCHASTIC

→many species have dumped it! Pufferfish / carvivorous bladderwart 97% coding / birds and bats 50% DNA (smaller cells*, more energy efficient, lighter) / coelacanth – extinct 63 mya – low transposons & static for 400 my – illustrates the RISK [*tyranosourous and velocorapters].

AGAIN It's not the junk per se that's beneficial, it's the process that created it and the side effects of it that are beneficial.

CONCLUSION: 53-90% is truly junk…

Introns

SPLICING – FIGURE 2.5

Example one – stopped the making of mRNA entirely, like jamming a photocopier.

Example two – expansion gets copied, acts like sponge mopping up proteins that bind to it and are no longer available to other RNAs to get processed, transported, survival time, efficiency of encoding proteins. Myotonic dystrophy.

Shows that introns are critical in some ways for regulating how mRNAs are used by the cells, a level of control, fine tuning.

Regulatory DNA – enhancers and loops

Promotor switch. Relocation of promotors → cancer

Enhancers = graduated response. Where/how much? Can't identify by sequence (explain), and not always functional, depends on previous activity/history -- memory makes easier to switch on again. Many enhancers for each gene; specific activation pattern for each cell lineage and state. Hard to tell which gene it effects because not nearby. Fuzzier. ALBERTS FIGURE.

RECENT eRNA (specific class of lncRNAs) as molecular traps that recruit TFs – via Mediator protein complex, RNA binds it and directs it to gene, protein modifies histones.

In embryonic stem cells, super-enhancers.

--LOOPS – accessable DNA is like pasta in cooking water. Additional Cohesin complex. Such long range interactions not restricted to enhancers. These happen a lot. Complex net of interactions, much flexibility in "regulating the tapestry of gene expression."

--LOOPS level 2 – to reach FACTORIES – DNA reeled through it, even from different chromosomes. Factories copy more than one RNA at a time – related functions. Factory has multiple parallel assembly lines, then assembles these parts into final product. eg hemoglobin.

ncRNA

Extra copies of chromosomes are devastating. Why aren't females crippled by their extra X chromosome? Inactivated. One of them switched off in each cell. Twisting a five meter towel until it's one millimeter. (Barr body) One X expresses a gene Xist. Non-coding RNA, doesn't leave nucleus, sticks to X and spreads along it, painting it. Nobody knows how, after 20 years of research. It binds proteins that shut down expression. [mosaic protects against hemophilia and color blindness, some muscular dystrophy]

Long ncRNA – over 10,000 types. Expressed as low level so weren't detected until recently. Hard to identify because can't match with sequenced animals, predictable patterns. Also not very conserved = meaningless or rapid evolution. Plus 3D structure might be conserved. [fig 8.1 if asked] Bystander events, byproducts, like wood chips from log cutting. Hard to test for functions when you don't have a hypothesized effect to look for. Many are in introns.

Knock down experiments in **150 intergenic**. Caused differentiation. 90% controlled expression of proteins either directly or indirectly, sometimes hundreds at a time. Usually genes far away. Expression of 75% dropped as cells moved from pluripotent to committed. So most of them are keeping the cells in embryonic state. Cf HOX gene region is rich in ncRNAs, and low in ancient virus fossils. Decreasing these experimentally causes limb deformities. Also relate to different types of cancer. Most cancer variants are intergenic.

Most in brain. Many specific to human or primate. Synaptogenesis. Development of human cortex. Alzheirmers. Neuropathic pain ion channel.

How they work: epigenetics

At, on, in addition to. CpG methylation is very stable, on/off. More flexible is histone modification, graded expression, the volume control—because different histones and 60 different modifier groups. ALBERTS FIGURE.

Enzymes that add these are blind to sequence. How do they know where to go? ANSWER: Other ncRNAs acting like the enhancers do, attracting histone modifying proteins—repressor, silencing. The ncRNA is transcribed nearby the target gene, moored there. Act nonspecifically, just depends on where they are—directors of the silencing patterns = changes in differentiation = easy metazoan evolution.

Repressors recognize pattern of modifications already there and reinforce them. Or other modifications will promote gene expression and inhibit the repressors. Opposite effect where ncRNAs promote expression by stopping methylating enzymes. Many layers. Many cancers are when repressors are overactive, wrong genes. "Delicate balance of lncRNAs and epigenetic modifiers, and disturbing the equilibrium is dangerous for the cell or individual."

EPIGENETIC CONTROL OF GENE EXPRESSION MIGHT HAVE ORIGINATED TO KEEP THE TRANSPOSONS UNDER CONTROL! Ditto with splicing, a second way ERV triggered multicellular life.

NO MAN'S LAND

Epigenetic modifcations are blind to sequence, so can creep into adjacent genes. THUS segments with no histones. Difference cell times need to insulate different regions. Determined by complex interactions between genome and proteins expressed at given time – via a supercomplex protein "pianist." Also by having the many tRNA genes in the insulating regions, activity blocking.

**

“The data indicate that the process of copying DNA into RNA, called transcription, is fundamentally noisy. The transcription factors that tell the cell where to start making RNAs are rather promiscuous, often having affinities for DNA sequences that will appear at random every 1,000 bases. They also act against a background where the packaging of DNA in the cell, which determines their access to such sequences, is dynamically changing. Apparently, wherever these changing conditions allow, transcription will start.

“This suggest that regulatory elements around genes act less to specifically start transcription there and more to make gene transcription at the gene more probable than the general background of RNA noise. Other examples of extra transcribed bases result from a transcription stopping mechanism that also appears to be noisy. In several cases, RNAs were found that started in one gene and plowed straight through to the next one down the chromosome.

“--That said, the study also found clear signals of defined transcription start sites at a rate of nearly ten times the number of genes in the area, suggesting some aspects of the excess transcription are nonrandom.”

GLOSSARY

Code and sequence.

Conserved sequence

Differentiation (pluripotent, development)

Histone proteins

Enzyme

Gene, exon, intron

Virus, retrovirus, endogenous retrovirus, HERV